

RI1: Repository Information for

Šizling, A.L., P. Keil, E. Tjørve, K.M.C. Tjørve, J.D. Žárský, and D. Storch. 202X. Mathematically and biologically consistent framework for presence-absence indices of diversity.

Content:

TABLE R1: Definitions and Scaling Properties	2-6
THESES T1-T21: Theses and Evidences	7-14
FIGURE R2: Average Values	15
FIGURE R3: Interrelated Assemblages	16
FIGURE R4: Multisite Version of Equation 2 in Šizling <i>et al.</i> , (202X)	17
R5: i-Independence of and s-Dependence on Species Richness	18
R6: Data Description	19

*For the list of the references see Šizling *et al.* (202X)*

TABLE R1 (definitions and scaling properties): The table shows the definitions of diversity indices and their mutual scaling properties. The references (e.g., T01.4) attributed to the equation marks refer to the derivations of the relationships (see Theses and Proofs T01 for derivations). The equation marks without references are derived through simple rearrangements of the equations listed in the table. All listed scaling properties are exact except those marked as ‘ \cong ’. Where possible, indices are defined according to Koleff *et al.* (2003) and Gaston *et al.* (2007); see the column ‘Notes’ for exceptions. Consistent with Koleff *et al.* (2003), and Gaston *et al.* (2007), a refers to the number of species shared by the focal assemblages ($a \stackrel{\text{def}}{=} S_{X \cap Y}$), and b and c represent the numbers of species unique to the first and second assemblages ($b \stackrel{\text{def}}{=} S_X - S_{X \cap Y}$ and $c \stackrel{\text{def}}{=} S_Y - S_{X \cap Y}$). Further references can be found in the ‘Notes’ column.

Notation	Definition	scaling properties	Notes
β_{Wh0}	$\stackrel{\text{def}}{=} \frac{S_T}{\langle S_1 \rangle}$	Gaston <i>et al.</i> (2007) call β_{Wh0} as β_{Wh2} , for the case of two assemblages; β_{Wh2} scales with indices of Jaccard index family only if the two sub-assemblages are adjacent, but it has fundamentally different inferences. $\beta_{Wh2} \stackrel{(T01.4)}{=} 2 \frac{1}{1+J}$	Overall Whittaker (1960) beta-diversity of two sub-assemblages of one assemblage. (in this study), $S_T = S_{X \cup Y} = S_X + S_Y - S_{X \cap Y}$ and $\langle S_1 \rangle = (S_X + S_Y)/2$ Beta-diversity <i>sensu stricto</i> (T08)
J	$\stackrel{\text{def}}{=} \frac{a}{a+b+c}$ $\stackrel{\text{def}}{=} \frac{S_{X \cap Y}}{S_X + S_Y - S_{X \cap Y}}$	$J \stackrel{(T01.1)}{=} \frac{\beta'_{Sim}}{1 + R^{-1} - \beta'_{Sim}}$	Jaccard index (1912) reversed turnover or co-occurrence if applied on between two species level (T02)
β_{SR}	$\stackrel{\text{def}}{=} \frac{a+b+c}{2a+b+c}$ $\stackrel{\text{def}}{=} \frac{S_X + S_Y - S_{X \cap Y}}{S_X + S_Y}$	$\beta_{SR} \stackrel{(T01.2)}{=} \frac{1}{1+J}$	Schluter & Ricklefs (1993) turnover or reversed co-occurrence applied on between two species level (T02)
$\beta_{S\emptyset r}$	$\stackrel{\text{def}}{=} \frac{2a}{2a+b+c}$ $\stackrel{\text{def}}{=} \frac{2S_{X \cap Y}}{S_X + S_Y}$	$\beta_{S\emptyset r} \stackrel{(T01.3)}{=} 2 \frac{J}{1+J}$	Sørensen (1948) similarity, reversed turnover or co-occurrence applied on between two species level (T02)
β_{DICE}	$\stackrel{\text{def}}{=} \frac{2a}{b+c}$ $\stackrel{\text{def}}{=} \frac{2S_{X \cap Y}}{S_X + S_Y - 2S_{X \cap Y}}$	$\beta_{DICE} \stackrel{(T01.17)}{=} 2 \frac{J}{1-J}$	Raup & Crick (1979) reversed turnover or co-occurrence applied on between two species level (T02)

β_{BC}	$\stackrel{\text{def}}{=} 1 - \frac{a}{2a + b + c}$ $\stackrel{\text{def}}{=} 1 - \frac{S_{X \cap Y}}{S_X + S_Y}$	$\beta_{BC} =_{(T01.16)} \frac{1}{1 + J}$	Bray & Curtis (1957) simplified for incidence data Turnover or reversed co-occurrence if applied on between species level (T02)
β_{HK}	$\stackrel{\text{def}}{=} 1 - \frac{2a}{2a + b + c}$ $\stackrel{\text{def}}{=} 1 - \frac{2S_{X \cap Y}}{S_X + S_Y}$	$\beta_{HK} =_{(T01.5)} \frac{1 - J}{1 + J}$	Harte & Kinzig (1997) Turnover or reversed co-occurrence applied on between two species level (T02)
β_t	$\stackrel{\text{def}}{=} \frac{b + c}{2a + b + c}$ $\stackrel{\text{def}}{=} \frac{S_X + S_Y - 2S_{X \cap Y}}{S_X + S_Y}$	$\beta_t =_{(T01.6)} \frac{1 - J}{1 + J}$	Wilson & Shmida (1984) in Gaston <i>et al.</i> (2007) Baselga (2010a,2012) calls this index Sørensen dissimilarity Turnover or reversed co-occurrence applied on between two species level (T02)
β_G	$\stackrel{\text{def}}{=} \frac{b + c}{a + b + c}$ $\stackrel{\text{def}}{=} \frac{S_X + S_Y - 2S_{X \cap Y}}{S_X + S_Y - S_{X \cap Y}}$	$\beta_G =_{(T01.7)} 1 - J$ $\beta_G =_{(T01.6, T01.7)} 2 \frac{\beta_t}{1 + \beta_t}$	Gaston <i>et al.</i> (2001) Turnover or reversed co-occurrence applied on between two species level (T02)
β_{Sim}	$\stackrel{\text{def}}{=} \frac{\min(b, c)}{a + \min(b, c)}$ $\stackrel{\text{def}}{=} \frac{\min(S_X, S_Y) - S_{X \cap Y}}{\min(S_X, S_Y)}$	$\beta_{Sim} =_{(T01.8)} 1 - \beta'_{Sim}$	Simpson beta Simpson (1943); Lennon <i>et al.</i> (2001) reversed nestedness (T05)
β'_{Sim} (rN)	$\stackrel{\text{def}}{=} \frac{a}{a + \min(b, c)}$ $\stackrel{\text{def}}{=} \frac{S_{X \cap Y}}{\min(S_X, S_Y)}$	$\beta'_{Sim} =_{(T01.8)} 1 - \beta_{Sim}$ $\beta'_{Sim} =_{(T01.9)} \frac{1+R}{R} \frac{J}{1+J}$	Lennon <i>et al.</i> (2001), Šizling <i>et al.</i> (2016) β'_{Sim} equals rN (range nestedness) Šizling <i>et al.</i> (2009) if computed between two species ranges Nestedness (T05)
β_{jtu}	$\stackrel{\text{def}}{=} 2 \frac{\min(b, c)}{a + 2 \min(b, c)}$ $\stackrel{\text{def}}{=} 2 \frac{\min(S_X, S_Y) - S_{X \cap Y}}{2 \min(S_X, S_Y) - S_{X \cap Y}}$	$\beta_{jtu} =_{(T01.21)} 2 \frac{\beta_{Sim}}{1 + \beta_{Sim}}$	Introduced as turnover component of Jaccard dissimilarity in Baselga (2012) reversed nestedness (T05)

β_{gl}	$\stackrel{\text{def}}{=} 2 \frac{ b - c }{2a + b + c}$ $\stackrel{\text{def}}{=} 2 \frac{ S_X - S_Y }{S_X + S_Y}$	$\frac{R}{1 + R} \stackrel{(T01.10)}{=} \frac{2 - \beta_{gl}}{4}$ $\beta_{gl} \stackrel{(*)}{=} 2 \frac{1 - R}{1 + R}$	<p>Lennon <i>et al.</i> (2001) species richness contrast (T06) (* follows from the above T19.10 equation)</p>
R	$\stackrel{\text{def}}{=} \frac{a + \min(b, c)}{a + \max(b, c)}$ $\stackrel{\text{def}}{=} \frac{\min(S_X, S_Y)}{\max(S_X, S_Y)}$	$R \stackrel{(T01.10)}{=} \frac{2 - \beta_{gl}}{2 + \beta_{gl}}$ $R \stackrel{(T01.9)}{=} \frac{J}{\beta'_{Sim}(1 + J) - J}$	<p>Newbold <i>et al.</i> (2016) scales in one-to-one manner with β_{gl} by Lennon <i>et al.</i> (2001); reversed species richness contrast or species richness uniformity (T06)</p>
$\beta_{sne} \stackrel{(*)}{=} \beta_{NES}$	$\stackrel{\text{def}}{=} \beta_t - \beta_{Sim}$	<p>does not scale with other indices</p>	<p>Baselga (2012) (* it is labeled as β_{NES} in Baselga 2010) No family under examination (T09)</p>
β_{jne}	$\stackrel{\text{def}}{=} \beta_G - \beta_{jtu}$	<p>does not scale with other indices</p>	<p>Baselga (2012) No family under examination (T13)</p>
β_{rps}	$\stackrel{\text{def}}{=} 2 \frac{\min(b, c)}{a + b + c}$ $\stackrel{\text{def}}{=} 2 \frac{\min(S_X, S_Y) - S_{X \cap Y}}{S_X + S_Y - S_{X \cap Y}}$	<p>does not scale with other indices</p>	<p>Introduced as (one-for-one) replacement and labeled R_{rel} in Podani & Schmera (2011) no phenomenon in consideration</p>
β_{nps}	$\stackrel{\text{def}}{=} \frac{a + b - c }{a + b + c} \text{ if } a \neq 0 \text{ \& } b \neq c$ $\stackrel{\text{def}}{=} \frac{S_{X \cap Y} + S_X - S_Y }{S_X + S_Y - S_{X \cap Y}} \text{ if } S_{X \cap Y} \neq 0 \text{ \& } S_X \neq S_Y$ $\stackrel{\text{def}}{=} 0 \text{ if } S_{X \cap Y} = 0 \text{ or } S_X = S_Y$	$\beta_{nps} \stackrel{(T01.22)}{=} 1 - \beta_{rps}$ <p>if there were not for the extra conditions that make the index value zero</p>	<p>Introduced as nestedness and labeled N'_{rel} in Podani & Schmera (2011) no phenomenon in consideration</p>
β_{rib}	$\stackrel{\text{def}}{=} \frac{a}{a + c}$ $\stackrel{\text{def}}{=} \frac{S_{X \cap Y}}{S_Y}$	<p>$\beta_{rib} = \beta'_{Sim}$ where $S_Y = \min(S_X, S_Y)$ else it does not scale with other indices</p>	<p>Ruggiero <i>et al.</i> (1998) Nestedness or reversed turnover depending on the direction; S_Y stands for species richness of the approached assemblage if we leave assemblage X (T11)</p>

$\langle \beta_{Sim} \rangle$	$\stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=1}^n \beta_{Sim,i}$	$\langle \beta_{Sim} \rangle =_{(T01.11)} 1 - \langle \beta'_{Sim} \rangle;$ $\langle \beta_{Sim} \rangle \cong_{(T01.12)} D$	arithmetic mean across n Simpson indices as introduced in Lennon <i>et al.</i> (2001) (a pair wise index); it is labeled simply Sim in literature (e.g., Gaston <i>et al.</i> 2007) The scaling (12) works for only assemblages with a high number of sites (≥ 10) reversed Nestedness (T05,T03)
$\langle \beta'_{Sim} \rangle$	$\stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=1}^n \beta'_{Sim,i}$	$\langle \beta'_{Sim} \rangle =_{(T01.11)} 1 - \langle \beta_{Sim} \rangle;$ $\langle \beta'_{Sim} \rangle \cong_{(T01.11,T19.12)} 1 - D;$ $\langle \frac{J}{1+J} \rangle \cong_{(T01.13*)} \langle \beta'_{Sim} \rangle \langle \frac{R}{1+R} \rangle$	The scaling between $\langle \beta_{Sim} \rangle$ and D works only for assemblages with a high number of sites (≥ 10) Nestedness (T05,T03) (*) if $cov(\langle \beta'_{Sim} \rangle, \langle \frac{R}{1+R} \rangle) \cong 0$, which is likely
D	Discrepancy (defined with an algorithm)	$D \cong_{(T01.12)} \langle \beta_{Sim} \rangle =_{(T01.11)} 1 - \langle \beta'_{Sim} \rangle$	Brualdi & Sanderson (1999) Scaling (T01.12) works only for assemblages with a high number of sites (≥ 10) Reversed Nestedness (T03)
N_{NODF}	$\stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=1}^n N_{NODF,i}$ $N_{NODF,i} \stackrel{\text{def}}{=} 1 - \beta_{Sim,i}$ if the two assemblages vary in their species richness; otherwise $N_{NODF,i} \stackrel{\text{def}}{=} 0$.	does not scale with other indices it would scale with β'_{Sim} if there was not for condition $S_X = S_Y \Rightarrow N_{NODF} = 0$	Almeida-Neto <i>et al.</i> (2008) No family under examination (T12)
N_0	Number of gaps (defined with an algorithm)	does not scale with other indices	Patterson & Atmar (1986) For two assemblages (in this study), $N_0 \stackrel{\text{def}}{=} \frac{\min(b,c)}{2a+b+c}$ if the two assemblages vary in their species richness; otherwise $N_0 \stackrel{\text{def}}{=} 0$. No family under examination (T10)

$$N_C \quad (\text{defined with an algorithm}) \quad N_{C1} \stackrel{(T01.19)}{=} 1 - \beta_{SR} \stackrel{(T01.20)}{=} \frac{\beta_{Sgr}}{2} \stackrel{(T01.3)}{=} \frac{J}{1+J}$$

$$N_{C2} \xrightarrow{\langle S \rangle \rightarrow \infty} \stackrel{(T04)}{=} 2\beta'_{Sim} - 1$$

Wright & Reeves (1992),
 For two assemblages (in this study),
 $N_{C1} \stackrel{\text{def}}{=} \frac{S_{X \cap Y}}{S_X + S_Y}$ (standardized as in Gotelli and
 McCabe 2002); $N_{C2} \stackrel{\text{def}}{=} \frac{S_{X \cap Y} - E(S_{X \cap Y})}{\max(S_{X \cap Y}) - E(S_{X \cap Y})}$
 (standardized as in Wright & Reeves 1992)
 N_{C1} , reversed turnover (T02)
 N_{C2} Nestedness (T04)

Theses (theses and proofs): Theses T1-21 state whether or not each index from Table R1 satisfies the constraints of the spatial phenomena (Figure 3 in Šizling et al. 202X), the conditions for i-independence of the indices, and the scaling between indices within a family and between families. The evidences for the theses employ three parameters: a , (the number of shared species, $S_{X \cap Y}$), b (the number of species exclusive to the first assemblage, $S_X - S_{X \cap Y}$), and c (the number of species exclusive to the second assemblage, $S_Y - S_{X \cap Y}$). The individual arrangements in Figure 3 in Šizling et al. (202X), are characterized as follows: $a = 0$ for the arrangements r3 and r4; $b = 0$ in arrangements r1,r2,r5 and r6; $b = 0$ and $c = 0$ in arrangement r5; and by $c_{r6} < c_{r2} < c_{r1}$ in arrangements r1,r2,r6.

Understanding the evidence for the theses requires a basic knowledge of linear algebra. Specifically, one should know: what a system of linear equations is and how it can be converted to a matrix (e.g., system: $a_{11}x_1 + a_{12}x_2 = b_1$; $a_{21}x_1 + a_{22}x_2 = b_2$; the matrix: $\begin{bmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \end{bmatrix}$); the Gauss elimination method; Cramer's rule and the theorem stating that a nonzero determinant indicates a set of mutually independent (in our work, i-independent) equations while a zero determinant indicates a set of mutually (i-)dependent equations. Formore details on the link between linear algebra and indices of diversity see Box 2 and Box 3 in Šizling et al. (202X).

T01 (scaling properties): *Relationships between the focal indices obey the equations listed in Table R1.* The evidence for the relationships is as follows (the numbering '1-22' refers to the labels associated with equation marks in Table R1; brackets $\langle . \rangle$ denote a mean value):

1. $J = \frac{a}{c+b+a} = \frac{a}{(c+a)+(b+a)-a} = \frac{\frac{a}{\min[b,c]+a}}{\frac{c+a}{\min[b,c]+a} + \frac{b+a}{\min[b,c]+a} - \frac{a}{\min[b,c]+a}} = \frac{\beta'_{Sim}}{1+R^{-1}-\beta'_{Sim}}$.
2. $\beta_{Sr} = \frac{a+b+c}{2a+b+c} = \left(1 + \frac{a}{a+b+c}\right)^{-1} = \frac{1}{1+J}$.
3. $\beta_{S\emptyset r} = \frac{2a}{2a+b+c} = 2 \left(1 + \frac{a+b+c}{a}\right)^{-1} = 2(1+J^{-1})^{-1} = 2 \frac{J}{J+1}$.
4. $\beta_{Wh2} = \frac{a+b+c}{(2a+b+c)/2} \stackrel{(T19.2)}{=} 2\beta_{Sr} = 2 \frac{1}{1+J}$.
5. $\beta_{HK} = 1 - \frac{2a}{2a+b+c} \stackrel{(T19.3)}{=} 1 - \beta_{Sor} = 1 - 2 \frac{J}{J+1} = \frac{1-J}{1+J}$.
6. $\beta_t = \frac{b+c}{2a+b+c} = \frac{2a+b+c-2a}{2a+b+c} = 1 - \frac{2a}{2a+b+c} \stackrel{(T19.5)}{=} \frac{1-J}{1+J}$.
7. $\beta_G = \frac{b+c}{a+b+c} = \frac{a+b+c-a}{a+b+c} = 1 - \frac{a}{a+b+c} = 1 - J$.
8. $\beta_{Sim} = \frac{\min[b,c]}{a+\min[b,c]} = \frac{a+\min[b,c]-a}{a+\min[b,c]} = 1 - \frac{a}{a+\min[b,c]} = 1 - \beta'_{Sim}$.
9. $JR(1 + R^{-1} - \beta'_{Sim}) \stackrel{(T01.1)}{=} \beta'_{Sim} \Rightarrow J(R + 1 - \beta'_{Sim} \cdot R) = \beta'_{Sim} \cdot R \Rightarrow \frac{J}{1+J} = \beta'_{Sim} \frac{R}{1+R}$.
10. Let $b \leq c$ then $R = \frac{b+a}{c+a}$ and thus $\beta_{gl} = 2 \frac{c-b}{2a+b+c} = 2 \frac{(c+a)-(b+a)}{2a+b+c} = 2 \frac{(c+a)(1-R)}{2a+b+c} = 2(1-R) \left(\frac{2a+b+c}{a+c}\right)^{-1} = 2(1-R) \left(\frac{a+b+a+c}{a+c}\right)^{-1} = 2(1-R)(1+R)^{-1} = 2 \frac{1-R}{R+1} \Rightarrow R = \frac{2-\beta_{gl}}{2+\beta_{gl}}$.
11. $\beta_{Sim} \stackrel{(T01.8)}{=} 1 - \beta'_{Sim} \Rightarrow \langle \beta_{Sim} \rangle = \langle 1 - \beta'_{Sim} \rangle \Rightarrow \langle \beta_{Sim} \rangle = \langle 1 \rangle - \langle \beta'_{Sim} \rangle$
12. Simulation based evidence (see Figure R2 below).
13. $\frac{J}{J+1} \stackrel{(T01.9)}{=} \beta'_{Sim} \frac{R}{1+R} \Rightarrow \langle \frac{J}{1+J} \rangle = \langle \beta'_{Sim} \frac{R}{1+R} \rangle \Rightarrow \langle \frac{J}{1+J} \rangle = \langle \beta'_{Sim} \rangle \langle \frac{R}{1+R} \rangle + \text{cov}(\beta'_{Sim}, \frac{R}{1+R})$.

Where the i-th values $\beta'_{Sim,i}$ and $\frac{R_i}{1+R_i}$ are s-independent across all i, $\text{cov}(\beta'_{Sim}, \frac{R}{1+R}) = 0$.

Nestedness, $\beta'_{Sim,i}$ and species richness uniformity $\frac{R_i}{1+R_i}$ does not constraint each other (share only one sr component of the $S_X, S_Y, S_{X \cap Y}$) and are likely s-independent (implying zero covariance). However, an ecological driver makes covariance nonzero and then the relationship in Table R1 does not work.

14. ($0 \leq R \leq 1 \Rightarrow 0 \leq \frac{R}{1+R} \leq 0.5$), it follows ($0 \leq \frac{R}{1+R} \leq 0.5 \ \& \ \frac{J}{1+J} =_{(T01.9)} \beta'_{Sim} \frac{R}{1+R} \Rightarrow \frac{J}{1+J} \leq 0.5\beta'_{Sim} \Rightarrow J \leq \frac{\beta'_{Sim}}{2-\beta'_{Sim}}$).

15. $J =_{(T01.6)} \frac{1-\beta_t}{1+\beta_t} =_{(Def \beta_{sne})} \frac{1-\beta_{sne}-\beta_{Sim}}{1+\beta_{sne}+\beta_{Sim}}$.

16. $\beta_{BC} \stackrel{def}{=} 1 - \frac{S_{X \cap Y}}{S_X + S_Y} = 1 - J \frac{S_X + S_Y - S_{X \cap Y}}{S_X + S_Y} = 1 - J\beta_{BC} \Rightarrow \beta_{BC} = 1 - J\beta_{BC} \Rightarrow \beta_{BC} = \frac{1}{1+J}$.

17. $\beta_{DICE} \stackrel{def}{=} \frac{2S_{X \cap Y}}{S_X + S_Y - 2S_{X \cap Y}} = 2J \frac{S_X + S_Y - S_{X \cap Y}}{S_X + S_Y - 2S_{X \cap Y}} = 2J \left(1 - \frac{S_{X \cap Y}}{S_X + S_Y - S_{X \cap Y}}\right)^{-1} = 2 \frac{J}{1-J}$.

18. $\beta_{sne} \stackrel{def}{=} \beta_t - \beta_{Sim} = \frac{S_X + S_Y - 2S_{X \cap Y}}{S_X + S_Y} - \frac{S_{X \cap Y}}{S_Y} = \frac{S_X S_{X \cap Y} - S_Y S_{X \cap Y}}{S_Y(S_X + S_Y)} = \frac{S_Y S_{X \cap Y} - S_Y S_{X \cap Y}}{S_Y(S_X + S_Y)} = \frac{(R^{-1}-1)S_{X \cap Y}}{S_X + S_Y}$.

$S_Y \leq S_X$ is an arbitrary choice without losing generality.

19. $N_{C1} \stackrel{def}{=} \frac{S_{X \cap Y}}{S_X + S_Y} = \frac{S_X + S_Y + S_{X \cap Y} - S_X - S_Y}{S_X + S_Y} = 1 - \frac{S_X + S_Y - S_{X \cap Y}}{S_X + S_Y} = 1 - \beta_{SR}$.

20. $N_{C1} \stackrel{def}{=} \frac{S_{X \cap Y}}{S_X + S_Y} = \frac{1}{2} \frac{2S_{X \cap Y}}{S_X + S_Y} = \frac{\beta_{S\sigma r}}{2}$.

21. $\beta_{jtu} \stackrel{def}{=} 2 \frac{\min(S_X, S_Y) - S_{X \cap Y}}{2 \min(S_X, S_Y) - S_{X \cap Y}} = 2 \frac{1-\beta'_{Sim}}{2-\beta'_{Sim}} = 2 \frac{\beta_{Sim}}{1+\beta_{Sim}}$.

22. $\beta_{nps} + \beta_{rps} \stackrel{def}{=} \frac{S_{X \cap Y} + |S_X - S_Y|}{S_X + S_Y - S_{X \cap Y}} + 2 \frac{\min(S_X, S_Y) - S_{X \cap Y}}{S_X + S_Y - S_{X \cap Y}} =_{(S_Y \leq S_X)} \frac{S_{X \cap Y} + S_X - S_Y}{S_X + S_Y - S_{X \cap Y}} + 2 \frac{S_Y - S_{X \cap Y}}{S_X + S_Y - S_{X \cap Y}} =$

$\frac{S_X + S_Y - S_{X \cap Y}}{S_X + S_Y - S_{X \cap Y}} = 1$. (*)=no discontinuity conditions

T02: Jaccard index (J) is consistent with the constraints on reversed turnover.

Note: All the indices that follow strictly increasing function of J are consistent with the constraints on reversed turnover, and all indices that follow a strictly decreasing function of J are consistent with the constraints on turnover. For the scaling, see Fig. 5 in Šizling et al. (202X), Thesis T01, and Table R1. For a new index, consult Repository RI4: Calculator.

Evidence: $J = \frac{a}{a+b+c}$. Here $a = 0$ in re-arrangements r3 and r4 in Figure 3 in Šizling et al. (202X).

Hence $J[r3] = J[r4] = 0$, which excludes the species richness contrast but supports turnover as defined in our framework. Additionally, $J[r1] < J[r2] < J[r6] < J[r5]$ in Figure 3 in Šizling et al. (202X). Hence, $0 < J[r1] < J[r5] = 1$, which supports only the reversed constraints of turnover.

T03: Discrepancy (D) for a high number of sites approaches a function that is consistent with the constraints on nestedness.

Note: Figure R2 shows that more than 10 sites are sufficient to achieve consistency with the constraints.

Evidence: For a small number of sites to compare, index D does not follow any of the focal phenomena. This evidence is based on D computed for various pairs of sets ($M1 = M2 = 15$, $M3=10$, $M4 = M5 = M6 = M7 = M8 = M9 = M10 = M11=5$, and $M12=4$). The rough (non-standardized) Discrepancy is defined as the minimum number of incidences that must be shifted along rows of an incidence matrix (rows represent sites; columns represent species) to achieve absolutely nested assemblages (Brualdi and Sanderson 1999). This definition simplifies to ‘ $\min(S_X, S_Y) - S_{X \cap Y}$ ’ for two assemblages. Index D is standardized by the number of incidences within the focal matrix, thus $D = \min(S_X, S_Y) - S_{X \cap Y} / (S_X + S_Y)$ in the case. The computed order is therefore $0 = D[r1] = D[r2] = D[r5] = D[r6] < D[r3] < D[r4] = 0.5$. The equality $D[r1] = D[r5]$ excludes species richness contrast, nestedness, and turnover, and the inequality $D[r3] < D[r4]$ excludes nestedness, but not contrast. However, if the number of sets is large enough (simulations suggest more than 30 simulations), the standardization by the number of incidences begins to work properly and the index D will scale with the indices of nestedness (Figure R2 below). We demonstrate the reason using two extreme cases: with maximum and minimum possible D . The Discrepancy of an absolutely nested matrix is by definition zero (no shift of incidences is needed). Maximum Discrepancy occurs in a matrix where almost all incidences (except the incidences of one site) must be shifted to gain an absolutely nested matrix. The Discrepancy of an absolutely non-nested matrix (where each site has its unique set of species) is then computed as the total species richness (sum across all sites) of species that

has to be shifted (i.e., excluding one of the most species rich sites). This equals $\sum_{i=1}^{Sites} S_i - \max_{i=1}^{Sites} [S_i]$, where S_i is the species richness of the i -th site. It is standardized by the number of incidencies, i.e., $\sum_{i=1}^{Sites} S_i$. Hence, $D = \frac{\sum_{i=1}^{Sites} S_i - \max_{i=1}^{Sites} [S_i]}{\sum_{i=1}^{Sites} S_i} = 1 - \frac{\max_{i=1}^{Sites} [S_i]}{\sum_{i=1}^{Sites} S_i}$, which approaches one if $\max_{i=1}^{Sites} [S_i] \ll \sum_{i=1}^{Sites} S_i$. This is the case of practically all datasets with large numbers of sites, and therefore $D \cong 1$ can be attributed to absolutely non-nested multisite assemblages. It is apparent that all matrices between these two extremes have D values between 0 and 1. The condition $\max_{i=1}^{Sites} [S_i] \ll \sum_{i=1}^{Sites} S_i$ might only be violated if the maximum species richness was high and species richness of the other sites was extremely small, which is unlikely. Simulations show a one-to-one scaling of D with Simpson beta (non nestedness) for matrices of 30 and 100 sites (Figure R2).

T04: N_C , standardized as in Wright & Reeves (1992) (here labeled N_{C2} ; labeled as C in Wright & Reeves 1992) approaches i -dependence on Simpson beta as Species Richness approaches infinity. N_{C2} is i -dependent on species richness.

Note: N_{C2} is thus consistent with the constraints on Nestedness.

Evidence: $N_{C2} \stackrel{\text{def}}{=} \frac{S_{X \cap Y} - E(S_{X \cap Y})}{\max(S_{X \cap Y}) - E(S_{X \cap Y})}$, where $\max(S_{X \cap Y}) = S_Y$ if we put $S_X \geq S_Y$, and $E()$ stands for expectation. So, $E(S_{X \cap Y}) = \frac{1 + \dots + S_Y}{S_Y} = \frac{S_Y + 1}{2}$. Hence $N_{C2} = \frac{2S_{X \cap Y} - S_Y - 1}{S_Y - 1}$. In general, $N_{C2} = \frac{2rS_{X \cap Y} - rS_Y - 1}{rS_Y - 1}$, where $r > 0$ emulates variation in species richness. Apparently, N_{C2} is i -dependent on r , and $\lim_{r \rightarrow \infty} N_C = 2 \frac{S_{X \cap Y}}{S_Y} - 1 = 2\beta'_{Sim} - 1$.

T05: Simpson beta β_{Sim} and β_{jtu} are consistent with the constraints on reversed nestedness and β'_{Sim} is consistent with the constraints on nestedness.

Note: The evidence is done for β_{Sim} . For the strictly increasing function of β_{jtu} with β_{Sim} , and the strictly decreasing function of β'_{Sim} with β_{Sim} see Fig. 5 in Šizling et al. (202X), Thesis T01, and Table R1.

Evidence: $\beta_{Sim} = \frac{\min[b,c]}{a + \min[b,c]}$. $\beta_{Sim}[r3] = \beta_{Sim}[r4] = 1$ because $a = 0$ in these cases; $\beta_{Sim}[r1] = \beta_{Sim}[r2] = \beta_{Sim}[r5] = \beta_{Sim}[r6] = 0$ because one of the variables b, c equals zero and $a \neq 0$ in these cases. All the indices are continuous and none of them depends on $\max(b, c)$. For scaling between β_{jtu} , β'_{Sim} and β_{Sim} see (T01.8, T01.21 and Figure 5 in Šizling et al. 202X).

T06: β_{gl} and R are consistent with the constraints on species richness contrast and species richness uniformity (i.e., reversed species richness contrast), respectively.

Evidence: R is the ratio of minimum to maximum species richness. Hence $R[r1] = R[r3] < R[r2] < R[r6] < R[r4] = R[r5] = 1$. R is therefore inversely related to species richness contrast, representing species richness uniformity. β_{gl} scales negatively with R , (see T01.10).

T07: β_{nps} is inconsistent with the constraints on nestedness, turnover or species richness contrast.

Evidence: If we ignore the extra conditions that make β_{nps} discontinuous (i.e., β_{nps} if $S_{X \cap Y} = 0$ or $S_X = S_Y$) then $0 = \beta_{nps}[r4] < \beta_{nps}[r3] < \beta_{nps}[r5] = \beta_{nps}[r6] = \beta_{nps}[r2] = \beta_{nps}[r1] = 1$, where $r1, \dots, r6$ are re-arrangements of two assemblages defined in Figure 3 in Šizling et al. (202X). The inequality $\beta_{nps}[r4] < \beta_{nps}[r3]$ excludes nestedness, the equality $\beta_{nps}[r2] = \beta_{nps}[r1]$ excludes the species richness contrast, and the equality $\beta_{nps}[r2] = \beta_{nps}[r1]$ excludes turnover.

If we add the condition: " $\beta_{nps} = 0$ if $S_{X \cap Y} = 0$ " (the variant of the index under discussion in Šizling et al. 202X, which is labeled N_{rel} in Podani and Schmera 2011), then the relations become $0 = \beta_{nps}[r4] = \beta_{nps}[r3] < \beta_{nps}[r5] = \beta_{nps}[r6] = \beta_{nps}[r2] = \beta_{nps}[r1] = 1$. This again excludes species richness contrast and turnover, but it agrees with the constraints on nestedness (Figure 3 in Šizling et al. 202X). This is only due to the extra condition, so if we

consider that indices must follow the requirement of continuity, the β_{nps} does not capture the nestedness. However, there is another reason why β_{nps} does not capture the experience of nestedness. The constraints $N[r1] = N[r2]$ implies independence of nestedness from the species richness of the richer assemblage. If we violate this condition where there is an overlap between the assemblages (i.e., $S_{X \cap Y} > 0$), then we are inconsistent with the constraint $N[r1] = N[r2]$ (Figure 3 in Šizling et al. 202X). The intuition is that if we enter the sea, then the fraction of how much we are immersed does not depend on the size of the sea. It is easy to see that $\beta_{nps} \stackrel{\text{def}}{=} 2(S_Y - S_{X \cap Y}) / (S_X + S_Y - S_{X \cap Y})$ changes with changing $S_X (= \max(S_X, S_Y))$, which excludes nestedness (arbitrarily without loss of universality $S_Y \leq S_X$).

If we add the condition: " $\beta_{nps} = 0$ if $S_{X \cap Y} = 0$ or $S_X = S_Y$ " (this variant of the index that is labeled N'_{rel} in Podani and Schmera 2011) then the relations turn into $0 = \beta_{nps}[r4] = \beta_{nps}[r3] = \beta_{nps}[r5] = \beta_{nps}[r6] = \beta_{nps}[r2] = \beta_{nps}[r1] = 1$. Again, the equality $\beta_{nps}[r3] = \beta_{nps}[r5]$ excludes nestedness, the equality $\beta_{nps}[r2] = \beta_{nps}[r1]$ excludes the species richness contrast, and the equality $\beta_{nps}[r2] = \beta_{nps}[r1]$ excludes turnover.

T08: Index β_{rps} is inconsistent with the constraints on nestedness, turnover and species richness contrast. The index β_{rps} is also inconsistent with the concept of One-for-One Replacement.

Evidence: The relations for β_{rps} are reversed to the relations for β_{nps} where no extra conditions that make the index discontinuous are considered (i.e., $0 = \beta_{rps}[r4] < \beta_{rps}[r3] < \beta_{rps}[r5] = \beta_{rps}[r6] = \beta_{rps}[r2] = \beta_{rps}[r1] = 1$), for $\beta_{rps} = 1 - \beta_{nps}$ in this case (T01.22). Therefore, the reasoning for nestedness, species-richness contrast, and turnover follows the evidence T07 above. In the case of one-for-one replacement, there is a constraint $R[r4] = R[r3]$ on the one-for-one replacement that follows the point of Podani & Schmera (2011). The inequality $\beta_{rps}[r4] < \beta_{rps}[r3]$ therefore excludes one-for-one replacement. Because the numerator of β_{rps} for r3 equals the numerator of β_{rps} for r4, the problem is introduced by denominator (i.e., standardization of the index with $S_X + S_Y - S_{X \cap Y}$). One may argue that the one-for-one replacement of r3 differs from one-for-one replacement of r4, and thus the inequality $\beta_{rps}[r4] < \beta_{rps}[r3]$ does not disqualify β_{rps} as a measure of the one-for-one replacement. In this case, however, there should be a difference between $\beta_{nps}(S_X = 8, S_Y = 20, S_{X \cap Y} = 1)$ and $\beta_{nps}(17, 20, 10)$, but $\beta_{nps}(8, 20, 1) = \beta_{nps}(17, 20, 10)$ (Figure 1 in Šizling et al. 202X).

T09: Nestedness by Baselga (β_{sne}) is inconsistent with the constraints on nestedness, turnover and species richness contrast.

Evidence: $\beta_{sne} = \frac{b+c}{2a+b+c} - \frac{\min(b,c)}{a+\min(b,c)}$. For re-arrangements r1 and r2 in Figure 3 in Šizling et al. (202X), $\beta_{sne}[r1] < \beta_{sne}[r2]$, because $\beta_{sim}[r1] = \beta_{sim}[r2] = 0$, $b = 0$, $c_{r2} < c_{r1}$, and a does not vary between these two cases. This excludes nestedness. $\beta_{sne}[r5] = \beta_{sne}[r3] = 0$, because $b = c$, which excludes turnover. $\beta_{sne}[r1] > \beta_{sne}[r3] = 0$, which excludes species richness contrast ($\beta_{sne}[r3] = 0$, because $a = 0$; and $\beta_{sne}[r1] > 0$, because $\min(b, c) = 0$ and $b \neq c$).

T10: Number of gaps (N_0) is inconsistent with the constraints on nestedness, turnover and species richness contrast.

Evidence: This evidence is based on N_0 indices computed for various pairs of sets (M1 = M2 = 15, M3=10, M4 = M5 = M6 = M7 = M8 = M9 = M10 =M11=5, and M12=4.). The computed order is $0 = N_0[r1] = N_0[r2] = N_0[r4] = N_0[r5] = N_0[r6] < N_0[r3] = 0.3$ for the index that was standardized by number of incidences within the focal matrix. This matches no experience of the five spatial phenomena.

T11: The Ruggiero index of beta-diversity (β_{Rib}) is consistent with the constraints on reversed turnover or nestedness depending on the direction.

Evidence: The index β_{Rib} depends on the order of the focal assemblages. It is defined as $\beta_{Rib} = \frac{a}{a+c}$, where c captures either the first or second assemblage. In our case, $0 = c_{r5} < c_{r6} <$

$c_{r2} < c_{r1}$ and $a_{r5} = a_{r6} = a_{r2} = a_{r1}$. Hence, $0 < \beta_{Rib}[r1] < \beta_{Rib}[r5]$. At the same time $0 = \beta_{Rib}[r3] < \beta_{Rib}[r4]$, because $a = 0$ in these cases. The index β_{Rib} thus captures reversed turnover. If we replace b with c , then $\beta_{Rib}[r1] = \beta_{Rib}[r2] = \beta_{Rib}[r6] = \beta_{Rib}[r5] = 1$, because $c=0$ in these cases. The index rib therefore captures nestedness in the case.

T12: *Nestedness by Almeida-Neto et al. (2008) (N_{NODF}) would be consistent with the constraints on nestedness if we ignored the condition that $S_X = S_Y \implies N_{NODF} = 0$.*

Evidence: Where assemblages differ in their species richness, N_{NODF} equals β'_{Sim} from the Simpson beta family, which is an index of nestedness. Where assemblages have equal species richness $N_{NODF} = 0$. If we accepted that two equal sized assemblages are not mutually nested ($NODF(r5) = 0$, but $NODF(r6) = 1$), then N_{NODF} would be an index of nestedness. However, our framework excludes this possibility.

T13: *The Nestedness resultant component of Jaccard dissimilarity (β_{jne}) is inconsistent with the constraints on nestedness, turnover or species richness contrast.*

Evidence: This evidence is based on N_0 indices computed for various pairs of sets ($M1 = M2 = 15$, $M3=10$, $M4 = M5 = M6 = M7 = M8 = M9 = M10 =M11=5$, and $M12=4$.). The computed order is $\beta_{jne}[r3] = \beta_{jne}[r4] = \beta_{jne}[r5] = 0$, $\beta_{jne}[r6] = 0.2$, $\beta_{jne}[r2] = 0.5$, and $\beta_{jne}[r1] \cong 0.7$. The equality $\beta_{jne}[r4] = \beta_{jne}[r5]$ excludes nestedness and turnover, and the equality $\beta_{jne}[r3] = \beta_{jne}[r4]$ excludes species richness contrast.

T14: *An algorithm to compute Jaccard similarity and Simpson nestedness from two i-independent indices if one of the indices is partitioned.*

The algorithm: Let

$$I_1 = \frac{K_{1a}S_{X \cap Y} + K_{2a}S_X + K_{3a}S_Y}{L_{1a}S_{X \cap Y} + L_{2a}S_X + L_{3a}S_Y} + \frac{K_{1b}S_{X \cap Y} + K_{2b}S_X + K_{3b}S_Y}{L_{1b}S_{X \cap Y} + L_{2b}S_X + L_{3b}S_Y} = I_a + I_b$$

(if index I_b is subtracted the k_{ib} coefficients are multiplied by '-1')

and

$$I_2 = \frac{K_{12}S_{X \cap Y} + K_{22}S_X + K_{32}S_Y}{L_{12}S_{X \cap Y} + L_{22}S_X + L_{32}S_Y}.$$

Then the triple I_a, I_b , and I_2 are necessarily i-dependent. Thus

$$\det \begin{bmatrix} I_a L_{1a} - K_{1a} & I_a L_{2a} - K_{2a} & I_a L_{3a} - K_{3a} \\ I_b L_{1b} - K_{1b} & I_b L_{2b} - K_{2b} & I_b L_{3b} - K_{3b} \\ I_2 L_{12} - K_{12} & I_2 L_{22} - K_{22} & I_2 L_{32} - K_{32} \end{bmatrix} \equiv 0.$$

I_2 , and I_1 are known ($I_1 = I_a + I_b$), so

$$\det \begin{bmatrix} (I_1 - I_b)L_{1a} - K_{1a} & (I_1 - I_b)L_{2a} - K_{2a} & (I_1 - I_b)L_{3a} - K_{3a} \\ I_b L_{1b} - K_{1b} & I_b L_{2b} - K_{2b} & I_b L_{3b} - K_{3b} \\ M_1 & M_2 & M_3 \end{bmatrix} \equiv 0$$

where $M_i = I_2 L_{i2} - K_{i2}$.

After expansion

$$I_b^2 D_{LL} - I_b(I_1 D_{LL} - D_{KL} + D_{LK}) + I_1 D_{LK} - D_{KK} = 0,$$

$$\text{where } D_{LL} = \det \begin{bmatrix} L_{1a} & L_{2a} & L_{3a} \\ L_{1b} & L_{2b} & L_{3b} \\ M_1 & M_2 & M_3 \end{bmatrix}; D_{KK} = \det \begin{bmatrix} K_{1a} & K_{2a} & K_{3a} \\ K_{1b} & K_{2b} & K_{3b} \\ M_1 & M_2 & M_3 \end{bmatrix}; D_{LK} = \det \begin{bmatrix} L_{1a} & L_{2a} & L_{3a} \\ K_{1b} & K_{2b} & K_{3b} \\ M_1 & M_2 & M_3 \end{bmatrix};$$

$$\text{and } D_{KL} = \det \begin{bmatrix} K_{1a} & K_{2a} & K_{3a} \\ L_{1b} & L_{2b} & L_{3b} \\ M_1 & M_2 & M_3 \end{bmatrix}.$$

Having I_b , we can pick up two i-independent indices and then use the algorithm from Box 3 in Šizling et al. XXXX to compute any dimensionless index.

This solution works only if the indices I_a, I_b are mutually i-independent, the indices I_a, I_1 are mutually i-independent, and the indices I_1, I_b are mutually i-independent (together they are always i-dependent). If any pair of indices is mutually i-dependent, the solution is simpler: convert the i-dependent indices to an index from their family (Table R1) and then follow the algorithm from Box 3 in Šizling et al. XXXX.

T15: Symmetric and linear indices with equal denominators are i-dependent.

That is: Let us have two indices that are defined as

$$I_\alpha \stackrel{\text{def}}{=} \frac{k_{1\alpha}f_1(S_{X \cap Y}) + k_{2\alpha}f_2(S_X) + k_{3\alpha}f_3(S_Y)}{l_{1\alpha}f_1(S_{X \cap Y}) + l_{2\alpha}f_2(S_X) + l_{3\alpha}f_3(S_Y)}, \quad (k_{1\alpha-3\alpha}, l_{1-3} \in \mathbf{R}), |l_{2\alpha}| + |l_{3\alpha}| > 0 \quad (\text{T15.1})$$

$$I_\beta \stackrel{\text{def}}{=} \frac{k_{1\beta}f_1(S_{X \cap Y}) + k_{2\beta}f_2(S_X) + k_{3\beta}f_3(S_Y)}{l_{1\beta}f_1(S_{X \cap Y}) + l_{2\beta}f_2(S_X) + l_{3\beta}f_3(S_Y)}, \quad (k_{1\beta-3\beta}, l_{1-3} \in \mathbf{R}), |l_{2\beta}| + |l_{3\beta}| > 0 \quad (\text{T15.2})$$

where $f_i(X)$ are strictly increasing or strictly decreasing functions from \mathbf{R}^+ to \mathbf{R}^+ , and $k_{2\alpha} = k_{3\alpha}, k_{2\beta} = k_{3\beta}$, and $l_{1\alpha} = l_{2\alpha}, l_{1\beta} = l_{2\beta}$,

then $(l_{1\alpha} = l_{1\beta} \ \& \ l_{2\alpha} = l_{2\beta}) \Rightarrow I_\alpha$ is i-dependent on I_β .

Note 1: if $f_i(X) = X, \forall X \in \mathbf{R}$, then Eqs. T15.1 and T15.2 turn into Eq. 8 in Šizling et al. 202X.

Note 2: It follows that two asymmetric indices with equal denominators are i-independent.

Note 3: The symmetry required here is a stronger form of symmetry as it necessitates symmetry in both the numerator and denominator separately. For example, the index R is not a symmetric index according to this theorem.

Evidence:

$$\begin{aligned} \det \begin{bmatrix} I_\alpha l_1 - k_{1\alpha} & I_\alpha l_2 - k_{2\alpha} & I_\alpha l_3 - k_{3\alpha} \\ I_\beta l_1 - k_{1\beta} & I_\beta l_2 - k_{2\beta} & I_\beta l_3 - k_{3\beta} \\ 0 & 1 & 1 \end{bmatrix} \\ = (I_\alpha l_1 - k_{1\alpha})(I_\beta l_2 - k_{2\beta}) + (I_\alpha l_3 - k_{3\alpha})(I_\beta l_1 - k_{1\beta}) - (I_\alpha l_1 - k_{1\alpha})(I_\beta l_3 - k_{3\beta}) \\ - (I_\alpha l_2 - k_{2\alpha})(I_\beta l_1 - k_{1\beta}) \\ = I_\alpha (k_{1\beta}(l_2 - l_3) + l_1(k_{3\beta} - k_{2\beta})) + I_\beta (l_1(k_{2\alpha} - k_{3\alpha}) + k_{1\alpha}(l_3 - l_2)) \\ + k_{1\alpha}(k_{2\beta} - k_{3\beta}) + k_{1\beta}(k_{3\alpha} - k_{2\alpha}) \end{aligned}$$

Therefore determinant is zero (i.e., indices are i-dependent) if $l_2 - l_3 = 0 \ \& \ k_{3\alpha} - k_{2\alpha} = 0 \ \& \ k_{3\beta} - k_{2\beta} = 0$.

T16: (*species rich and poor assemblages*): The variation of $S_{poor} = 2\langle S \rangle R / (1 + R)$ indicates inevitable change in richness of the species poorer assemblage, and the variation of $S_{rich} = 2\langle S \rangle / (1 + R)$ indicates inevitable change in the richness of the species richer assemblage.

Evidence: The evidence is based on the solution of the system of three equations for J, β'_{sim} and $\langle S \rangle$. See Box 2 in Šizling et al. 202X. If S_X, S_Y , and $S_{X \cap Y}$ are mutually i-independent, the system

$$\text{follows } \begin{bmatrix} J & J & -J-1 & 0 \\ 0 & \beta'_{sim} & -1 & 0 \\ 1 & 1 & 0 & 2\langle S \rangle \end{bmatrix} \sim \begin{bmatrix} J & J & -J-1 & 0 \\ 0 & \beta'_{sim} & -1 & 0 \\ 0 & 0 & J+1 & 2J\langle S \rangle \end{bmatrix} \Rightarrow S_{X \cap Y} = 2\langle S \rangle \frac{J}{J+1} \ \& \ S_Y = \frac{S_{X \cap Y}}{\beta'_{sim}} \ \&$$

$$S_X = \frac{(J+1)S_{X \cap Y} - JS_Y}{J}. \text{ It follows that } S_Y = 2\langle S \rangle \frac{J}{(J+1)\beta'_{sim}} \stackrel{(Eq.2)}{=} 2\langle S \rangle \frac{R}{1+R}. \text{ The last equality follows}$$

$$\text{from Eq. 2. Finally, } S_X = \frac{2\langle S \rangle J - 2\langle S \rangle \frac{R}{R+1}}{J} = 2\langle S \rangle \left(1 - \frac{R}{R+1}\right) = \frac{2\langle S \rangle}{R+1} = \frac{S_Y}{R}. \text{ Because } 0 < R \leq 1, \text{ then}$$

$S_Y \leq S_X$, and we relabel $S_{poor} := S_Y$ and $S_{rich} := S_X$. If S_X, S_Y , and $S_{X \cap Y}$ are mutually i-dependent, then $\beta'_{sim} = 1$ or $R = 1$ and the derived equations for S_{poor} and S_{rich} remain valid.

T17: *Jaccard index (J) is consistent with conceptual predefinition of co-occurrence if applied to more sites occupied by two species instead of more species that belong to two assemblages.*

Evidence: If we consider the Venn diagrams in Figure 3 (Šizling et al. 202X) to show two species ranges (sets of occupied sites), then the constraints on Co-occurrence follow:

$$\text{Min} = \text{Co}[r4] = \text{Co}[r3] < \text{Co}[r1] < \text{Co}[r2] < \text{Co}[r6] < \text{Co}[r5] = \text{Max}.$$

Because $0 = J[r4] = J[r3] < J[r1] < J[r2] < J[r6] < J[r5] = 1$, the evidence is completed.

T18: The system of Eq. 10 and Eq. 11 in Šizling et al. (202X) is mutually i-independent if $S_{X \cap Y} \neq 0$ and $S_X \neq S_Y \neq S_{X \cap Y} \neq S_X$.

Evidence: If S_X, S_Y , and $S_{X \cap Y}$ are mutually i-independent, then the system follows

$$\begin{bmatrix} J & J & -J-1 & 0 \\ 0 & \beta'_{Sim} & -1 & 0 \\ 1 & 1 & 0 & 2\langle S \rangle \end{bmatrix} \sim \begin{bmatrix} J & J & -J-1 & 0 \\ 0 & \beta'_{Sim} & -1 & 0 \\ 0 & 0 & J+1 & 2J\langle S \rangle \end{bmatrix}. \text{ If } J \neq -1 \text{ and } J \neq 0 \text{ and } \langle S \rangle \neq 0 \text{ (mean}$$

S), this matrix provides a unique and non-trivial (i.e., non-zero) solution of the system.

Alternatively, $\det \begin{bmatrix} J & J & -J-1 \\ 0 & \beta'_{Sim} & -1 \\ 0 & 0 & 1+J \end{bmatrix} = J(1+J) \cdot \beta'_{Sim} \neq 0$, if $J, \beta'_{Sim} \neq 0$. Since $S_{X \cap Y} \neq 0$ implies

$J \neq 0, \beta'_{Sim} \neq 0$, and $\langle S \rangle \neq 0$, it follows that $2J\langle S \rangle \neq 0$ and the solution is not trivial (it is non zero). As an alternative, the last scaling equation can be replaced with $S_X + S_Y - S_{X \cap Y} = S_{TOT}$ where S_{TOT} is species richness of both sites. If the condition ' $S_{X \cap Y} \neq 0$ and $S_X \neq S_Y \neq S_{X \cap Y} \neq S_X$ ' is violated, then each index definition is a linear function of two variables and thus two equations are enough to get a unique solution, making one of the equations i-dependent on the others.

T19: (*i-dependence of three indices*): If $S_{X \cap Y} \neq 0$ (i.e., $J \neq 0$), then the value of any index defined by Eq. 8 in Šizling et al. (202X) can be computed from values of Jaccard index (J) and Simpson nestedness (β'_{Sim}) using

$$I = \frac{J(\beta'_{Sim}(k_1+k_2)-k_2+k_3)+\beta'_{Sim}k_2}{J(\beta'_{Sim}(l_1+l_2)-l_2+l_3)+\beta'_{Sim}l_2}. \quad \text{(Equation 12 in Šizling et al. 202X)}$$

Evidence: Consider arbitrary $S_Y \leq S_X$. Then $S_{X \cap Y} \neq 0 \Rightarrow S_Y \neq 0 \Rightarrow S_X \neq 0$. The new index is defined as $I = \frac{k_1 S_{X \cap Y} + k_2 S_X + k_3 S_Y}{l_1 S_{X \cap Y} + l_2 S_X + l_3 S_Y}$ ($|l_2| + |l_3| > 0$ & $|k_2| + |k_3| > 0$), which is the definition Eq. 8 in Šizling et al. (202X). The Eq. 12 results as a solution of the determinant equation

$$\det \begin{bmatrix} -J & -J & J+1 \\ 0 & \beta'_{Sim} & -1 \\ Il_2 - k_2 & Il_3 - k_3 & Il_1 - k_1 \end{bmatrix} = 0. \text{ The proof of the i-independence follows the reverse}$$

logic. (i) substitute definitions of J and β'_{Sim} into Eq. 12, (ii) convert Eq. 12 into the third equation of the system of linear equations and (iii) show that the determinant equals zero. In the same way is proved the equation for triplets I, J, R and I, β'_{Sim}, R . In these cases the restriction $S_{X \cap Y} \neq 0$ is replaced with $S_Y \neq 0$.

T20: If Šizling et al. (2016) published only two partitioned indices (β_{sne}, β_{nps}) then the referred information would be unequivocal.

Evidence: When we convert the indices reported by Šizling et al. (2016) into two partitioned indices, we got $\beta_{sne} = 0.1, \beta_{nps} = 0.61$ for the pre-agricultural landscape, and $\beta_{sne} = 0.00, \beta_{nps} = 0.51$ for early agricultural landscape. These indices suggest: $S_X \cong 31.01, S_Y \cong 22.99, S_{X \cap Y} \cong 15.48$ or $S_X \cong 41.63, S_Y \cong 12.37, S_{X \cap Y} \cong 2.28$ for pre-agricultural landscape; and $S_X \cong 32.00, S_Y \cong 32.00, S_{X \cap Y} \cong 21.60$ or $S_X \cong 48.32, S_Y \cong 15.68, S_{X \cap Y} \cong 0.00$ for early agricultural landscape (thesis T14 or Repository R4: calculator). This can be interpreted as either increasing species richness uniformity (R increases from 0.74 to 1), or decreasing species richness uniformity (R decreases from 0.74 to 0.32) species richness uniformity. This dual interpretation precludes any correct conclusion.

T21: *Any dimensionless index of turnover (defined by the constraints in Fig. 3) must have both S_X and S_Y in its denominator.*

Evidence: Dimensionless indices are ratios with numerators and denominators. Accurate measurement of turnover requires that information on both assemblages S_X , and S_Y be represented to capture their interaction fully. Including only one of S_X , or S_Y in the denominator while having both in the numerator creates an index without an upper limit. This contradicts the required bounded nature of the index, as specified by the constraints $T[r4] = T[r3] = \text{Max}$ in Fig. 3.

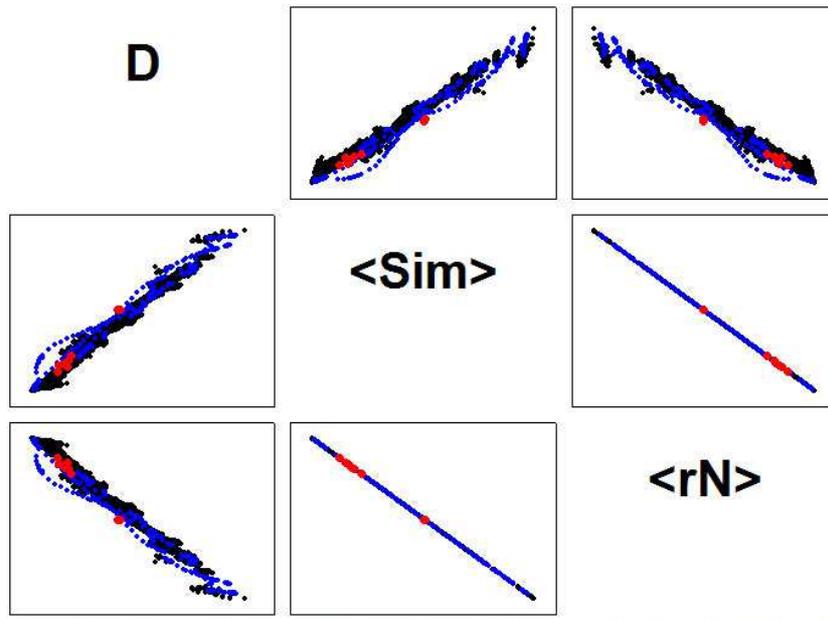


FIGURE R2 (average values): Relationships between indices of nestedness computed for more assemblages as simulated (30 sites x 30 species – black dots; 100 sites x 100 species – blue dots) and observed (red dots, R6 - data description). Matrices were generated to cover a wide spectrum of assemblages (see simulation details below). Indices include D (discrepancy by Brualdi & Sanderson, 1999); $\langle Sim \rangle$ (simple mean of the Simpson index across all pairs of adjacent plots; Lennon et al., 2001); and $\langle rN \rangle \equiv \langle \beta'_{Sim} \rangle$ (simple mean of nestedness as defined in Šizling et al., 2009); D , was standardized by dividing with the total number of incidences within the focal matrix (Greve et al., 2005). The indices scale with each other in a one-to-one manner, belonging to the same family and measuring the extent of nestedness. For exact evidence, see theses T03, and T05.

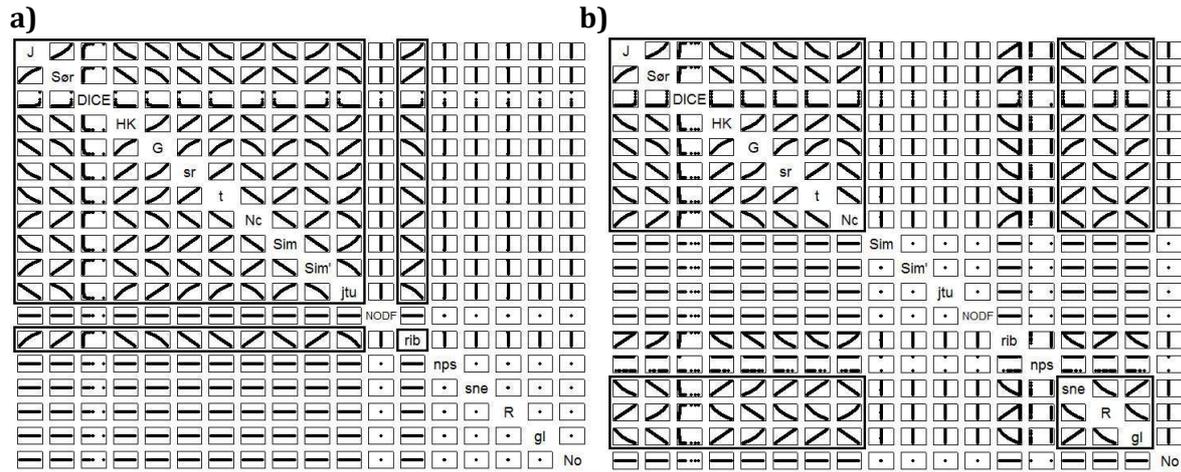


FIGURE R3 (interrelated assemblages): Relationship between pairwise indices where there is no contrast in species richness (a) and where assemblages are perfectly nested (b). As predicted, no contrast in species richness merges families of J , and β_{Sim} together, and perfect nestedness merges families of J , β_{sne} , and β_{gl} together. The evidence that β_{sne} shows variability where nestedness is perfect disqualifies this measure from being a proxy for nestedness component. Black rectangles delimit the merged families. For a detailed legend, see capture to Figure 5 in Šizling *et al.* 202X.

The random points for Fig. R3 and Figs 5,6 in Šizling *et al.* (202X) follow $S_{X \cap Y} = Rnd(1) * (1 - T)$, $S_X = S_{X \cap Y} + trunc(100 * (1 - N) * Rnd(1))/100$, $S_Y = Trunc(100 * (Rnd(1) * (min(1, S_X + C) - max(0, S_X - C)) + max(0, S_X - C)))/100$. $T, N, C \in \langle 0,1 \rangle$ where $T=1$ indicates no shared species richness; $T=0$ indicates all possible overlaps in species richness; $N=1$ indicates absolute nestedness; $N=0$ indicates that nestedness is not constrained; $C=1$ indicates unlimited species richness contrast; $C=0$ indicates no species richness contrast; $G=0.2$ indicates a maximum 20% variation in the difference between species richness of the sites. In particular, $T=0, N=0, C=1$ (Figure 5 in Šizling *et al.* 202X); $T=0, N=0, C=0.2$ (Figure 6); $T=0, N=0, C=0$ (Figure R3A); $T=0, N=1, C=1$ (Figure R3B in Šizling *et al.* 202X).

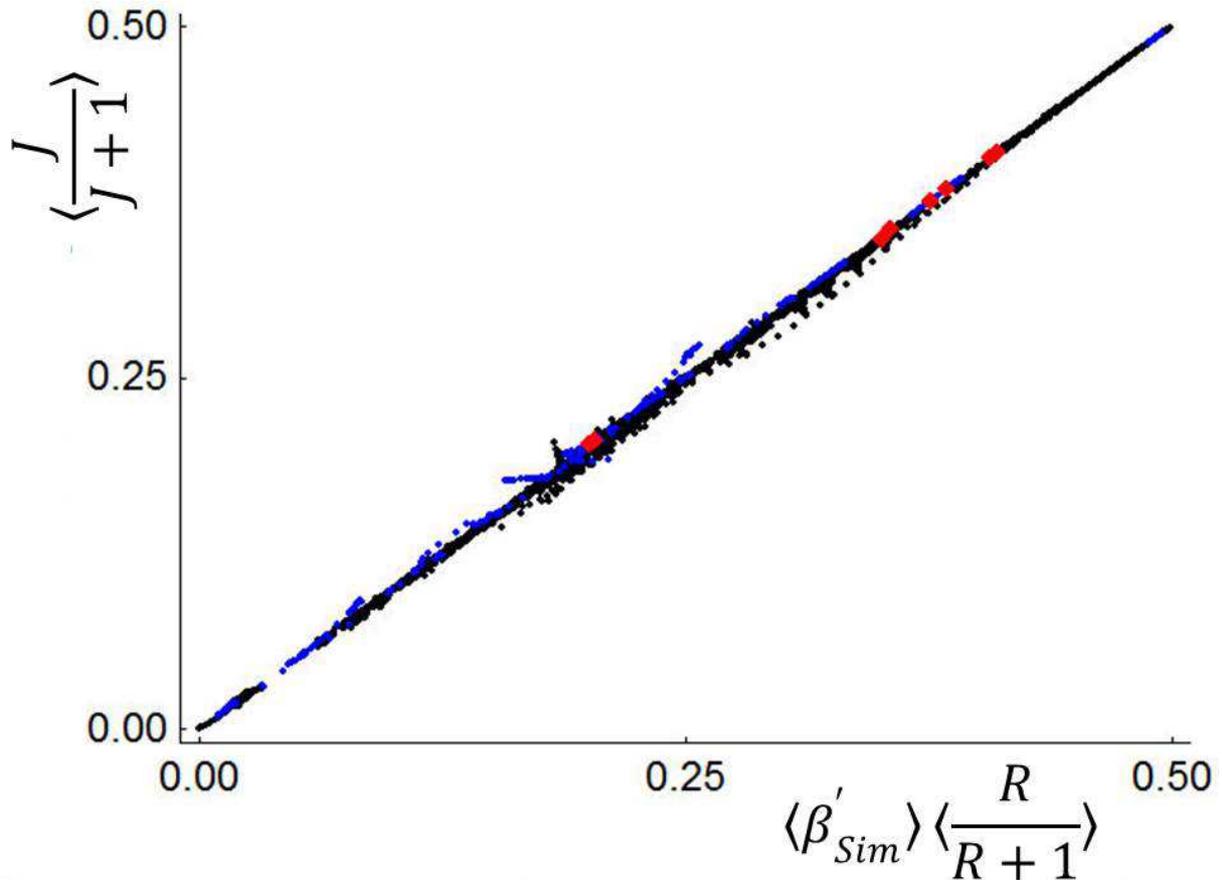


FIGURE R4 (Multisite Version of Eq. 2 in Šizling *et al.*, 202X): The accuracy of the approximate Eq. 2 applied to mean values across multiple pairs of sites follows thesis T01.13. The exact equality between the X-values and Y values in the plot is affected by the covariance between β'_{Sim} and $R/(1 + R)$, which is expected to be nearly zero. The figure shows the extent to which this expectation is met for simulated (30 sites x 30 species – black dots; 100 sites x 100 species – blue dots) and observed (red dots) multiple-site assemblages (R6 data description and RI2: Data). The y-axis shows the left side of the equation (i.e., the mean reversed turnover; $\langle J/(J + 1) \rangle$), and the x-axis shows the right side of the equation (i.e., the product of the mean values of nestedness, $\langle \beta'_{Sim} \rangle$ and species richness uniformity $\langle R/(R + 1) \rangle$). The points approach an identity line, showing that the covariance approaches zero as expected. The 30x30 and 100x100 assemblages were produced by the algorithm adopted from Šizling *et al.* (2009).

Šizling, A.L., E. Šizlingová, D. Storch, J. Reif and K. J. Gaston. 2009. Rarity, commonness, and the contribution of individual species to species richness patterns. *American Naturalist*, 174:82-93.

R5 (i-independence of and s-dependence on species richness):

The Jaccard index (J), and thus all the indices that scale one-to-one with J (Figure R5), has been referred to as ‘dependent’ on species richness (Simpson, 1943; Lennon *et al.*, 2001; Koleff & Gaston, 2002; Baselga, 2010a) and on the contrast between species richness of two assemblages (Simpson, 1943). This has led to a search for an index that is species richness ‘independent’, and to attempts to modify J , so that ecologists could compare assemblages that varied in species richness. This s-dependence between J and species richness was based on empirical experience (Koleff & Gaston, 2002), and on arguments that there are bounds on the J imposed by contrast in species richness (Simpson, 1943).

We found that all dimensionless (unitless) indices, including J , are i-independent of species richness. For all indices that can be expressed by the universal definition (Eq. 8 in Šizling *et al.* 202X) it holds that

$$I \stackrel{\text{def}}{=} \frac{k_1 S_{X \cap Y} + k_2 S_X + k_3 S_Y}{l_1 S_{X \cap Y} + l_2 S_X + l_3 S_Y} = \frac{k_1 \rho S_{X \cap Y} + k_2 \rho S_X + k_3 \rho S_Y}{l_1 \rho S_{X \cap Y} + l_2 \rho S_X + l_3 \rho S_Y}, (\rho \neq 0) \tag{Eq. R5}$$

and thus the index does not change when S_X , S_Y , and $S_{X \cap Y}$ scale proportionally to each other. In this case, the index has the same value regardless of species richness, and thus it is not uniquely determined by species richness. Any observed s-dependence between the dimensionless index and species richness is therefore caused by disproportional scaling between S_X , S_Y or $S_{X \cap Y}$.

Disproportional scaling can, however, appear at sites with small species richness because species richness is an integer. In this case, the frequency distribution of possible J values is affected by total species richness. This in turn affects the most likely value of J , imposing its s-dependence on species richness. The reason is that the Jaccard index can only have a finite number of values. For example, if $S_X = 1$ then $J = 1, 1/2, 1/3, 1/4, \dots, 0$, accumulating possible values below $1/2$. $S_X = 2$ then allows for $2/3$, which is above $1/2$, $S_X = 3$ allows for $3/4 > 2/3$ and so on. Further computation of possible J values for increasing S_X (Figure R5) shows an increasingly even distribution of J -values. This mechanisms works for any index that can be expressed by Eq. 8 in Šizling *et al.* (202X), and the effect cannot be eliminated by inventing a new dimensionless index.

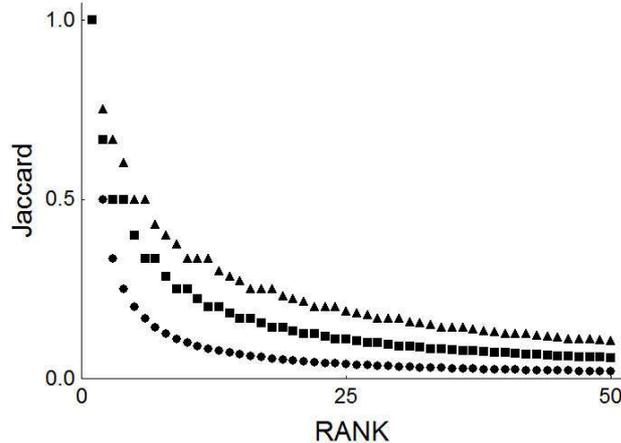


FIGURE R5: Rank plot of the first fifty values that can reach Jaccard index if S_X is fixed and S_Y and $S_{X \cap Y}$ vary within their limits ($1 \leq S_Y < \infty$ and $0 \leq S_{X \cap Y} \leq \min(S_X, S_Y)$); the three distributions on display correspond to $S_X = 1$ (circles), $S_X = 2$ (squares) and $S_X = 3$ (triangles).

R6 (data description):

For comparison with the results from the artificial (simulated) data, we also plotted observed values of the indices. This allows us to identify relationships that are mathematically possible but may be rare or absent in nature. These observations consisted of three different datasets: a set of 29 microbial assemblages extracted from cryoconite on the Greenland Ice Shield, a set of 24 arctic plant assemblages (4 from Greenland, and 20 from Svalbard), and a set of 20 temperate zone plant assemblages (10 from the Czech Republic, and 10 from Southern Norway). The microbial assemblages were sampled by J.Ž. and A.Š., and processed by J.Ž. Plant assemblages were sampled by A.L.Š., Eva Šizlingová and E.T (see Dataset RI.2).

A list of plant species found in a 10x10 m area was recorded. The data are nested in the sense that the assemblages are grouped so that each group of five assemblages is located within a 1km diameter circle. For the purposes of this analysis, only Genera were used.

Microbial assemblages were sampled at 300 m intervals along two lines on the western margin of the Greenland Ice Sheet in the vicinity of Kangerlussuaq. Sampling and sample processing procedures followed Cameron *et al.* 2016. Here we use data inferred from environmental RNA using Illumina amplicon sequencing to detect the active part of the microbial assemblage. Processing of the sequencing output was performed using the QIIME2 environment (Bolyen *et al.*, 2019), filtering for sequences present at least three times in the whole dataset, and rarefaction to the sampling depth of 5000 features per sample. This resulted in the exclusion of 7 samples out of 36 that had fewer features than the sampling depth. The remaining samples were used for the diversity analysis.

Cameron, K.A., M. Stibal, J. D. Žárský, E. Gozdereliler, M. Schostag and C. S. Jacobsen. 2016.

Supraglacial bacterial community structures vary across the Greenland ice sheet. *FEMS Microbiology Ecology* 92, doi: 10.1093/femsec/fiv164

Bolyen, E., J. R. Rideout, M. R. Dillon, *et al.* 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology* 37:852–857.